

Insulin, IGF-2 and Type 1 Diabetes Mellitus: Recently Implicated Genetic Loci

Cheri L. Deal, PhD, MD

*Department of Pediatrics, Ste-Justine Hospital,
University of Montreal
Montreal, Quebec, Canada*

Constantin Polychronakos, MD

*Department of Pediatrics
Montreal Children's Hospital, McGill University
Montreal, Quebec, Canada*

Recently, great strides have been made in elucidating the genetic components of insulin-dependent diabetes mellitus (IDDM), one of the best examples of a multifactorial disease with both environmental and polygenic etiologies. This article focuses on one of the more recently implicated and best investigated loci to date, *IDDM2*. While its effects are no doubt less than that of the major histocompatibility complex, human lymphocyte antigen (HLA), *IDDM2* maps to chromosome 11p15.5, where at least 2 candidate genes are found, those for insulin (*INS*) and for insulin-like growth factor 2 (*IGF2*).

Familial clustering and a high concordance rate in monozygotic twins indicate that genetic transmission of susceptibility is responsible for about half of the risk for the development of type 1 diabetes. The importance of the HLA locus, located on chromosome 6p21 (*IDDM1*), first came to light in the 1970s following association and linkage studies in affected and nonaffected siblings, and appears to account for 42% of the genetic component.¹ Type 1 diabetes is primarily a sporadic disease (90%); population-based (case-control) studies provided the early HLA association data. Family studies in which there were more than 1 affected child confirmed the association of specific HLA haplotypes (alleles) with type 1 diabetes (ie, the demonstration of linkage disequilibrium) and revealed preferential sharing of certain HLA haplotypes among affected sibs. The underlying biology of this linkage is not yet completely understood. It appears that the class II HLA molecules affect the immune response because of their highly polymorphic sequence variations, resulting in differences in the peptide-binding groove used in antigen presentation.²

That a locus involved in cellular immune recognition was involved in an autoimmune disease came as no surprise. However, 2 key questions remained: What other genes could play a role and how do we go about identifying these genes? Whereas reverse genetics (determination of the chromosomal location

and identification of new genes in spite of ignorance of the disease mechanism) has resulted in spectacular successes in identifying genes responsible for single-gene (mendelian) diseases, the application of this method to common polygenic phenotypes involves difficulties whose magnitude is hard to even estimate.³ To date, no *novel* gene has been identified and cloned on the basis of its linkage to a complex (multifactorial) disease phenotype. Linkage can only narrow the locus to within several centimorgans (cM). In human genetic maps, 1 cM roughly corresponds to 1 million bp of DNA (1 Mb) and contains, on average, about 20 genes. Furthermore, positional cloning of mendelian disorder genes relies on the occasional patient with a large deletion/insertion or chromosomal rearrangement to further narrow the disease locus; this is not an option in complex disorders like diabetes, in which disease susceptibility is encoded not by gene-inactivating mutations but by subtle DNA sequence variants common in the general population.

Global searches of the whole genome, using hundreds of equally spaced microsatellite markers to detect linkage to specific chromosome locations, help to orient our search for disease-related candidate genes based on our knowledge of their participation in particular biochemical or cellular pathways.^{4,5} While these studies are fraught with the statistical hazards of multiple comparisons and other methodologic controversies, they may help guide our quest for candidate genes. A disease is assigned to a chromosomal locus only after linkage has been formally demonstrated, replicated, and confirmed in at least 3 different datasets. The robustness of the linkage must be continually verified in additional datasets—preferably ones comprised of genetically homogeneous populations—and these regions need to be further saturated with markers to further define the loci. There are then association-based tests that take advantage of linkage disequilibrium (that is, the preferential association of specific marker polymorphisms with the disease) to narrow the locus to specific gene(s).

The presence of genes whose products are functionally related to the phenotype (candidate genes) in the region identified by linkage analysis can greatly accelerate this process. If no such genes exist, or if they are tried and ruled out, RNA sequences transcribed from the genetically defined interval (positional candidates) can be identified using transcript

maps of the human genome, such as the prototype recently published.⁶ After the genes are fully cloned, polymorphisms in and around them can be identified and examined for association with a specific disease, eg, for alleles that are more frequent in diabetic than in nondiabetic subjects. However, the sequence hypervariability seen at *IDDM1*—the class II HLA locus—is the evolutionary result of adaptation to a wide variety of pathogens and is unlikely to be found in other functionally significant proteins. In most cases, subtle coding or even noncoding variants have to be examined for biologic significance. Table 1 presents a summary of loci detected by genome-wide scans for type 1 diabetes susceptibility. Linkage disequilibrium studies for *IDDM2* narrowed the locus to a variable number of tandem repeat (VNTR) polymorphisms immediately upstream of the gene for insulin; this was facilitated by an intense and systematic search for markers flanking the *INS* gene on chromosome 11p15.5. The importance of this region is well known because of its involvement in diseases other than type 1 diabetes (many tumors, including Wilms' tumors and adrenocortical tumors, and Beckwith-Wiedemann syndrome; see previous reviews in *GGH* 1994;10(1):1-4,6-10. Unfortunately, the current human genome map, in general, does not afford this high degree of marker density.⁷⁻⁹

THE *IDDM2* LOCUS

Researchers suspected that insulin (acting as an antigen triggering autoimmune destruction of the pancreatic beta cell) may be the product of a candidate gene for type 1 diabetes, and linkage of this disease to a locus mapping to chromosome 11p15.5 (*IDDM2*) was established a decade ago (Figure 1). More recently, thorough association studies narrowed the locus to a polymorphism 356 bp upstream of the insulin gene (*INS*) promoter consisting of a VNTR of a 14-bp consensus sequence. A large number of alleles can be distinguished by size and by different variants of this consensus repeat unit. Most alleles in whites contain either 30 to 45 repeats (class I) or, less frequently, more than 150 repeats (class III). Intermediate (class II) alleles are rare.

Approximately 40% to 45% of whites have at least 1 class III allele, compared with <20% in diabetic patients. This suggests that class III (long) alleles are protective. This protective effect is observed in I/III heterozygotes and is, therefore, dominant. The relative risk for development of type 1 diabetes is increased 3- to 4-fold when a subject is homozygous for I/I (I/I versus I/III or III/III), and this specific polymorphism is estimated to account for 10% to 15% of the familial clustering of diabetes.¹⁰

Table 1
Type 1 Diabetes Susceptibility Loci in Humans*

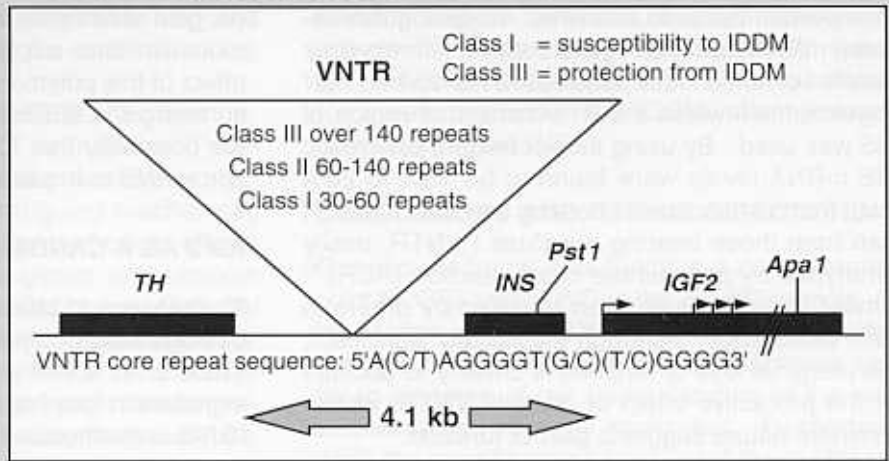
Locus	Chromosomal Location	Detected in Davies Genome-Wide Scan? ⁵	Detected in Other Datasets? [†]
<i>IDDM1</i> (HLA)	6p21	Yes	Yes
<i>IDDM2</i> (<i>INS</i> 5' VNTR)	11p15.5	Yes	Yes
<i>IDDM3</i>	15q26	Yes	Yes
<i>IDDM4</i>	11q13	Yes	Yes
<i>IDDM5</i>	6q25	Yes	Yes
<i>IDDM6</i>	18q21	Yes	Yes
<i>IDDM7</i>	2q31	No	Yes
<i>IDDM8</i>	6q25-q27	Yes	Yes
<i>IDDM9</i>	3q21-q25	No	Cited in Todd ⁴
<i>IDDM10</i>	10p11.2-q11.2	Yes	Yes
<i>IDDM11</i>	14q24.3-q31	No	Yes
<i>IDDM12</i> (<i>CTLA-4</i>)	2q33	No	Yes
<i>IDDM13</i>	2q34	No	Yes
<i>IDDM15</i> (distinct from HLA)	6p21	No	Yes
Not assigned (<i>GCK</i>)	7p	No	Yes
Not assigned	Xq	No	Cited in Todd ⁴

The *IDDM* nomenclature is officially assigned to a locus after linkage has been formally demonstrated, replicated, and confirmed in at least 3 independent datasets. [†]References to specific loci can be obtained by consulting *Online Mendelian Inheritance in Man* (<http://gdbwww.gdb.org/omim>).

Adapted with permission from Todd JA. *Proc Natl Acad Sci USA* 1995; 92:8560-8565.

Figure 1
The *IDDM2* Locus

The *IDDM2* locus, located on human chromosome 11p15.5, has been mapped to a variable number of tandem repeat sequences lying 3' to the tyrosine hydroxylase gene (*TH*) and 5' to the insulin gene (*INS*) and its close (less than 2 kb) neighbor, the gene for insulin-like growth factor 2 (*IGF2*), whose 4 promoters are indicated by the arrows. Useful RFLP exon polymorphisms to study allele-specific transcription are noted for *INS* (*Pst1*) and *IGF2* (*Apa1*). Note that the *IGF2* gene is expressed from the paternal gene copy in most tissues, whereas the other imprinted genes at the 11p15.5 locus (not shown) are expressed by the maternal gene copy (these include centromeric to *TH*: *p57KIP2* and *ASCL2*; telomeric to *IGF2*: *H19*).



FROM GENETICS TO GENE FUNCTION

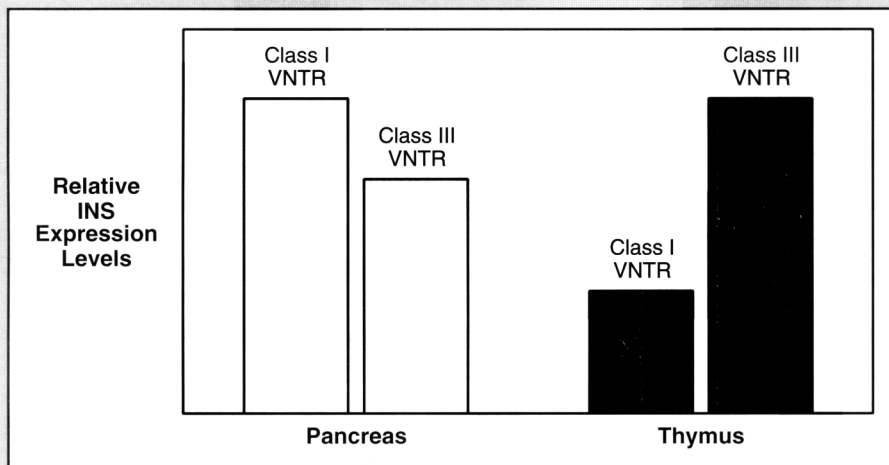
What biologic mechanism might account for this effect of the VNTR on diabetes susceptibility? Since the *INS* VNTR does not encode a protein sequence, it must exert transcriptional effects on nearby genes. Transcriptional effects of VNTRs elsewhere in the genome have been found, eg, those reported for *HRAS*¹¹ (Harvey rat sarcoma virus proto-oncogene), a membrane-associated small guanosine triphosphate (GTP)-binding protein believed to participate in the development and/or maintenance of certain malignancies. Although *INS* may be the principle target of its 5' VNTR, this VNTR also

could affect transcription of *IGF2*, the gene encoding insulin-like growth factor 2, whose promoters are 5 to 24 kb downstream (Figure 1), a distance compatible with enhancer effects. Therefore, elucidation of the *IDDM2* mechanism requires a systematic search for allelic effects on transcription levels of genes that are in close proximity to the VNTR in tissues of known importance in diabetes.

INS AS A CANDIDATE GENE FOR *IDDM2*

Recent studies have demonstrated that the VNTR does indeed exhibit tissue-specific allelic effects on *INS* transcription *in vivo* and *in vitro* (Figure 2).¹²⁻¹⁷

Figure 2



Schematic summarizing VNTR effects on *INS* mRNA levels in fetal tissue, genotyped for allele class. Data for fetal pancreas are adapted from Vafiadis et al.¹⁵; those for thymus are adapted from Vafiadis et al.¹⁶

Since the pancreas is the major tissue (and initially thought to be the only tissue) in which physiologically significant levels of insulin are produced, it was an important tissue to examine. To distinguish between mRNA from each gene copy in heterozygous samples of human fetal pancreas, a transcribed *Pst1* polymorphism within the 3' untranslated region of *INS* was used. By using this technique, pancreatic *INS* mRNA levels were found to be 15% to 20% lower from chromosomes bearing the class III VNTR than from those bearing the class I VNTR, easily genotyped by polymerase chain reaction (PCR).¹⁵ Similar findings have been reported by others in adult pancreas.¹² Although statistically significant, this marginal loss of function is unlikely to account for the protective effect of class III VNTR, whose dominant nature suggests gain of function.

Mouse thymus expresses insulin¹⁸ and many otherwise tissue-specific proteins (such as peptide hormones and exocrine pancreatic enzymes), presumably for the purpose of immune tolerance development. Human thymus also expresses insulin as mRNA and protein, albeit at levels 10³ to 10⁴ times lower than those in the pancreas. These observations prompted a search for *INS* VNTR allele-specific effects on thymic *INS* mRNA levels in 12 samples of human fetal thymus that were heterozygous at the VNTR (I/III) and informative for the transcribed *Pst1* polymorphism. Unlike pancreas, the thymic insulin mRNA value from VNTR class III haplotypes was 2- to 3-fold higher than from class I in 10 of the 12

samples ($P < 0.007$).¹⁶ Since antigen effects on thymocyte selection are dose-dependent,¹⁹ this increased *INS* expression from VNTR class III chromosomes (ie, gain of function) might enhance thymic tolerance to insulin, thus explaining the dominant protective effect of this polymorphism. This is in favor of insulin being the *IDDM2* gene, but does not rule out the possibility that *IGF2* is involved, either in addition to *INS* or in place of *INS*.

IGF2 AS A CANDIDATE GENE FOR IDDM2

The reasons to consider the product of *IGF2* in the pathogenesis of type 1 diabetes are summarized in Table 2. It is well appreciated that the IGF axis is important in lymphopoiesis and immune function.^{20,21} IGF-2 is synthesized in and has biologic effects on tissues that are important in the pathophysiology of diabetes, such as pancreas, thymus, and lymphocytes. It is a particularly important fetal mitogen that has been shown to stimulate proliferation and/or prevent apoptosis. It has been hypothesized that IGF-2 may either function as a tolerogenic autoantigen in the thymus²² or promote survival of self-reactive lymphocytes.²⁰ Involvement of IGF-2 in the pathogenesis of type 1 diabetes could explain parental imprinting effects reported in the genetics of *IDDM2*. Imprinting refers to differential transcriptional behavior (typically silencing) of a gene copy on the basis of the sex of the parent from whom it was inherited,²³ such as has been demonstrated for *IGF2* in some

Table 2
Arguments Supporting a Role for *IGF2* in Type 1 Diabetes Pathophysiology

Physiologic Arguments*

1. IGF-2 and type I IGF-1 receptor (mitogenic receptor) expressed in human fetal pancreas, thymus, and leukocytes
2. Type 2 IGF-2 receptor (involved in IGF-2 clearance) highly expressed in fetal thymus; suggests modulating IGF-2 levels may be important
3. IGF-2 transgenic animals show changes in lymphoid tissue, including clonal expansion of thymocytes, mainly mature CD4+ cells
4. IGF-2 known to promote cell survival (anti-apoptosis factor)

Genetic Arguments†

1. *INS* and *IGF2* mapped to within 2 kb of each other, a distance compatible with shared enhancer effects
2. Type 1 diabetes susceptibility in offspring of diabetic parents is related to paternal transmission (in some populations), suggesting imprinting effects
3. *INS* is not imprinted in human fetal pancreas; *INS* also is biallelically expressed in the majority of fetal thymi
4. *IGF2* is imprinted and expressed from the paternal allele in fetal pancreas and thymus
5. Placental *IGF2* mRNA levels are correlated with VNTR class
6. The 5' *INS* VNTR has transcriptional effects on *IGF2* (artificial constructs)

*See Polychronakos et al²⁰ for literature review

†References given in text.

tissues in the human fetus and in the human placenta.²⁴ In contrast, *INS* is not imprinted in human fetal or adult pancreas,^{12,20,25} despite the imprinted expression of the mouse insulin genes, *Ins1* and *Ins2*, in yolk sack, where only the paternal copies are transcribed.²⁶

Linkage²⁷ or association at *IDDM2* has been reported in alleles of paternal¹²⁸⁻³⁰ or maternal^{12,25} origin, but other studies have found no parent-of-origin effect.³¹ In addition to the intriguing mechanistic questions they raise, these parent-of-origin effects may provide clues as to the genes and biologic mechanisms involved at the *IDDM2* locus. Paternal bias would favor *IGF2* because of its exclusive paternal expression; an explanation for the maternal bias is less obvious, unless one postulates longer-range VNTR effects on imprinted, maternally expressed genes at the 11p15.5 locus.

We have recently demonstrated that higher steady-state *IGF2* mRNA levels are associated with paternal class I VNTR alleles in normal human placenta, a tissue in which *IGF2* is exclusively expressed from the paternal allele. Furthermore, by using a construct in which a class I or class III VNTR is placed upstream of an important fetal *IGF2* promoter (P3), greater reporter gene activity is observed with the class I VNTR allele. The physiologic significance of this could relate to the anti-apoptotic effect of IGF-2 on self-reactive T cells during fetal life, whereby the class I VNTR alleles, via increased *IGF2* transcription, favor the development of diabetes because of survival of self-reactive lymphocytes.³²

It should be stressed that dividing VNTR alleles into classes affords only a broad categorization, as each class contains many distinct alleles of varying sizes. It is possible that expression of *IGF2* and/or *INS* associated with particular VNTR alleles also is affected differently by parental imprinting, known to be a tissue-specific, developmentally-specific, and polymorphic phenomenon.²³ In the case of *INS*, the higher thymic transcript levels from the class III haplotype seen in 10/12 specimens is present independently of the parental origin of the class III. However, imprinting is likely involved in the complete silencing of the class III allele in the remaining 2/12 thymi.¹⁶ Pugliese et al¹⁷ also found monoallelic *INS* expression in 3/10 thymi. In all 5 cases, it was the class III haplotype that was silenced. Therefore, it is possible that imprinting of thymic *INS* in this minority of individuals requires the presence of specific class III alleles. A precedent for such haplotype-restricted imprinting has been described: the polymorphic silencing³³ of the paternal copy of *IGF2R* (the IGF-2 receptor gene) is controlled by a sequence variant in cis.³⁴ Imprinted *IGF2* expres-

sion might be modulated by a similar phenomenon in leukocytes, where there is variable "relaxation" of imprinting, thereby allowing transcription from the maternal copy to a variable extent in different individuals.^{15,20} Dependence of relaxation on specific VNTR alleles could explain the maternal effect, if *IGF2* were involved in the *IDDM2* effect instead of (or in addition to) *INS*.

CONCLUSIONS AND FUTURE DIRECTIONS

Genetic evidence and functional considerations point to *IGF2* and/or *INS* as the *IDDM2* gene(s). Our knowledge of how *IDDM1* and *IDDM2* contribute to the diabetes phenotype continues to increase and is far in advance of our understanding of the many additional loci recently implicated. Furthermore, while the specific genes involved at these other loci are unknown at present, it is of interest that several genes coding for products that interact with IGF-2 can be found within some of them, such as those coding for the type 1 (15q26) and type 2 (6q27) IGF receptors and the IGF-binding proteins (2q34, 7p13). Is this mere chance, or does it explain some of the epistatic genetic interactions, whereby these loci are not independent but may be acting on the same or overlapping pathways involved in the pathophysiology of diabetes? With the speed at which technologic advances have accelerated our understanding of diabetes in the past 5 years, one can hope to see these questions answered in the not too distant future.

REFERENCES

1. Owerbach D, et al. *Diabetes* 1996;45:544-551.
2. She J-X. *Immunol Today* 1996;17:323-329.
3. Lander ES, et al. *Science* 1994;265:2037-2048.
4. Todd JA. *Proc Natl Acad Sci USA* 1995;92:8560-8565.
5. Davies JL. *Nature* 1994;371:130-136.
6. Schuler GD, et al. *Science* 1996;274:540-546.
7. Elston RC. *Am J Hum Genet* 1997;60:255-262.
8. Thomson G. *Diabetes Reviews* 1997;5:106-115.
9. Schork NJ. *Diabetes Rev* 1997;5:116-122.
10. Bennett ST, et al. *Annu Rev Genet* 1996;30:371-403.
11. Green M, et al. *Genomics* 1993;17:429-434.
12. Bennett ST, et al. *Nat Genet* 1995;9:284-292.
13. Catignani Kennedy G, et al. *Nat Genet* 1995;9:293-298.
14. Lucassen AM, et al. *Hum Mol Genet* 1995;4:501-506.
15. Vafiadis P, et al. *J Autoimmun* 1996;9:397-403.
16. Vafiadis P, et al. *Nat Genet* 1997;15:289-292.
17. Pugliese A, et al. *Nat Genet* 1997;15:293-296.
18. Jolicoeur C, et al. *Proc Natl Acad Sci USA* 1994;91:6707-6711.
19. Ashton-Rickardt PG, et al. *Cell* 1994;76:651-663.
20. Polychronakos C, et al. *Develop Genet* 1995;17:253-262.
21. Clark R. *Endocr Rev* 1997;18:157-179.
22. Geenen V, et al. *Thymus* 1993;21:115-127.
23. Deal C. *Curr Opin Pediatr* 1995;7:445-458.
24. Giannoukakis N, et al. *Nat Genet* 1993;4:98-102.
25. Bennett et al. *J Autoimmun* 1996;9:415-421.
26. Giddings SJ, et al. *Nat Genet* 1994;6:310-313.
27. Julier C, et al. *Nature* 1991;354:155-159.
28. Polychronakos C, et al. *Diabetologia* 1995;38:715-719.
29. Pugliese A, et al. *J Autoimmun* 1994;7:687-694.
30. Bui MM. *J Autoimmun* 1996;9:97-103.
31. Bain SC, et al. *Nat Genet* 1992;2:212-215.
32. Paquette J, et al. *J Biol Chem* 1998. In press.
33. Xu YQ, et al. *Biochem Biophys Res Commun* 1993;197:747-754.
34. Xu YQ, et al. *Oncogene* 1997;14:1041-1046.